

VU Research Portal

Integrated Analytical Strategies for Drug Discovery

de Vlieger, J.S.B.

2011

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

de Vlieger, J. S. B. (2011). *Integrated Analytical Strategies for Drug Discovery*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Chapter 5

Determination and identification of estrogenic compounds generated with biosynthetic enzymes using hyphenated screening assays, high resolution mass spectrometry and off-line NMR.

Abstract

This paper describes the determination and identification of active and inactive estrogenic compounds produced by biosynthetic methods. A hyphenated screening assay towards the human estrogen receptor ligand binding domain (hER) α and hER β integrating target-ligand interactions and liquid chromatography - high resolution mass spectrometry was used. With this approach, information on both biologic activity and structure identity of compounds produced by bacterial mutants of cytochrome P450s was obtained in parallel. Initial structure identification was achieved by high resolution MS/MS, while for full structure determination, P450 incubations were scaled up and the produced entities were purified using preparative liquid chromatography with automated fraction collection. NMR spectroscopy was performed on all fractions for 3D structure analysis; this included 1D- ^1H , 2D-COSY, 2D-NOESY, and ^1H - ^{13}C -HSQC experiments. This multidimensional screening approach enabled the detection of low abundant biotransformation products which were not suitable for detection in either one of its single components. In total, the analytical scale biosynthesis produced over 85 compounds from 6 different starting templates. Inter and intra day variation of the biochemical signals in the dual receptor affinity detection system was less than 5 %. The multi-target screening approach combined with full structure characterization based on high resolution MS(/MS) and NMR spectroscopy demonstrated in this paper can generally be applied to e.g. metabolism studies and compound-library screening.

1. Introduction

Relatively new in early drug discovery is the use of biocatalysts capable of producing new chemical entities [1]. Besides the fact that these procedures are used for the production of drug metabolites and being part of the metabolite identification process [2-4], this approach is increasingly showing its potentials in the extension of the so-called chemical space. Cytochrome P450 BM3 is one of the most studied bacterial P450s and has the highest catalytic activity ever detected for a monooxygenase [5]. Although its natural substrates are long-chain fatty acids, several engineered mutants of BM3 are reported to convert drug and drug-like molecules [6-9]. Biosynthesis using the engineered BM3s can lead to the formation of multiple (active) products simultaneously, requiring advanced

analytical technologies for product identification and characterization. Conventional high throughput screening (HTS) assays lack the capability to fully structural characterize the hits in a mixture with multiple active compounds present. To individually assess the affinity of active compounds in mixtures towards selected drug targets, on-line high resolution screening (HRS) technologies were developed and recently applied to metabolism studies [10,11]. These receptor affinity detection (RAD) systems integrate liquid chromatography (LC) with affinity assessment and therefore allow the direct correlation between individual ligands and their affinity towards target proteins. Biochemical detection systems reported previously are based on fluorescence readout [12-14], fluorescent polarization [15] or mass spectrometry based reaction detection [16-20]. LC coupled to mass spectrometry (MS) is by far the most used tool in drug metabolism studies. The metabolic stability of drug candidates is typically monitored by a short LC analysis coupled to a triple-quadrupole mass spectrometer, enabling quantification of the metabolic conversion of drug candidates. For metabolite identification, high resolution instruments are commonly used and provide information such as accurate mass on the metabolites and their fragment ions. Applying *in vitro* metabolite screening in early drug discovery provides the medicinal chemists with information on the so-called soft spots in scaffold molecules [21]. The development of MS based strategies enabling the identification of drug metabolites is regularly reported [22,23]. The implementation of algorithms pre and post acquisition to predict metabolites [24], to filter mass defects [25], neutral losses and isotope patterns [26] significantly enhances the rate of success in metabolite identification. By combining the advantages of on-line screening and high resolution mass spectrometry (HR-MS/MS) for first line structure identification, multidimensional data is obtained, and a quick selection can be made based on affinity and the modification introduced by e.g. (bio)synthesis, metabolism or degradation. In the present paper, we describe the application of hyphenated screening assays, high resolution mass spectrometry and NMR for the determination and identification of biosynthetically produced estrogenic compounds. After the biosynthesis, all compounds generated are directly characterized with the on-line high resolution screening platform developed for this purpose to obtain information on structure and biologic activity in parallel. New biosynthetic products are characterized by a combination of HR-MS(/MS) and NMR. To demonstrate the strength of our

approach, we have selected the human estrogen receptor subtypes α and β (hER α and hER β) as model targets. These receptors belong to the nuclear receptor family and are involved in many hormonal regulation processes [27-29]. Currently these targets are used in cancer therapy, contraception and other hormone related therapies [30,31].

2. Experimental

2.1 Material and Methods

17 α -estradiol (17 α -E2), 17 β -estradiol (17 β -E2), 17 α -ethinylestradiol (EE2), estriol (E3), monosodium dihydrogen phosphate, disodium monohydrogen phosphate, glucose-6-phosphate, glucose-6-phosphate dehydrogenase, polyethyleneglycol 3350, dimethylsulfoxide-d₆, 99.9 atom %D (DMSO-d₆) and 3-(trimethylsilyl)propionic-2,2,3,3-d₄ acid, sodium salt, 98 atom %D (TSP) were purchased at Sigma-Aldrich (Schnelldorf, Germany). Maleic acid, 99%, was purchased from Acros Organics (Geel, Belgium). Applichem (Lokeren, Belgium) supplied β -nicotinamide adenine dinucleotide phosphate (NADPH) tetra sodium salt. Coumestrol was obtained from Fluka (Buchs, Germany). Sodium chloride was received from Riedel-de Haën (Seelze, Germany). Acetonitrile, methanol (ULC-MS grade) and formic acid were obtained from Biosolve (Valkenswaard, The Netherlands). Water was produced by a Milli-Q device of Millipore (Amsterdam, The Netherlands). Enzyme-linked immunosorbent assay (ELISA) blocking reagent was purchased from Roche Diagnostics (Penzberg, Germany). Norethisterone (NET), Org X and Org Y were obtained from the library of the Schering-Plough Research Institute (Oss, the Netherlands).

Human liver microsomes (HLM), pooled from 50 donors were obtained from Xenotech (Lenexa, USA) and contained 20 mg/mL protein (Lot No. 0710619). The bacterial P450 BM3 M11_{his} and M02_{his} mutants were prepared according to the protocol described recently [32]. The ligand binding domain (LBD) of the human Estrogen Receptor α was expressed and purified using the protocol described by Eiler *et al.* [33], without 17 β -E2 in the medium. Purified His-tagged LBD of the human Estrogen Receptor β was obtained from Schering-Plough Research Institute (Oss, the Netherlands).

2.2 Compound generation by combinatorial biosynthesis

2.2.1 Analytical Scale:

The selected compounds were subjected to biotransformation by several enzymatic systems. Biosynthetic incubations with the P450 BM3 mutants had a final volume of 1.5 mL and consisted of 250 nM enzyme, 100 μ M substrate in 100 mM potassium phosphate buffer pH 7.4 and were performed at 24 °C. The DMSO concentration was always below 2 %. A NADPH regenerating system was used to initiate the reactions, resulting in final concentrations of 0.2 mM NADPH, 0.3 mM glucose-6-phosphate, and 0.4 units/mL glucose-6-phosphate dehydrogenase. Biotransformation reactions with HLM were performed at 37 °C having a final volume of 1.5 mL and consisted of 1 mg/mL protein, 100 μ M substrate in 100 mM potassium phosphate buffer pH 7.4 containing 5 mM MgCl₂. At time points 0 min, 10 min and 60 minutes, 500 μ L samples were taken from the incubation and added to 1 mL of ice-cold methanol, vortexed for 30 seconds and centrifuged for 15 minutes at 6000 rpm. The supernatants were subsequently dried in a speedvac and reconstituted in 200 μ L of dried methanol. All samples were analyzed using the on-line LC–MS dual receptor affinity detection system.

2.2.2 Preparative scale

The metabolites originating from the biotransformation of NET were produced on a preparative scale using the bacterial P450 M02_{his} mutant. A 40 mL reaction volume containing 250 nM of the M02_{his} mutant, 500 μ M of NET and the NADPH regenerating system was made in 100 mM potassium phosphate buffer pH 7.4 at 24°C. After 3 hours of reaction, the products were extracted from the reaction mixture using three times 20 mL dichloromethane. The organic layers were subsequently collected and transferred into a round-bottom flask for evaporation using a rotary evaporator. The dried product was redissolved in 250 μ L DMSO and injected into a Waters preparative LC system, equipped with a Gilson Liquid Handler for injection and fractionation, a Waters Xbridge C18-MS (10 x 50 mm 3.5 μ m particles) column, and a Waters photodiode array detector. At a flow-rate of 5 mL/min the following gradient was used: 0-4 minutes 100% A1 (90% H₂O, 10%

acetonitrile); 4 – 22 minutes linear increase to 50 % B1 (10 % H₂O, 90 % acetonitrile), 22 – 24 minutes linear increase to 100 % B where kept constant for 2 minutes and finally a column equilibration time of 5 minutes with 100 % A. Fractions were collected triggered by UV absorbance at 243 nm. In addition, all other fractions were collected every minute to avoid the loss of samples showing no UV absorbance at the set wavelength. All fractions were dried under nitrogen and dissolved in 500 µL of DMSO-d₆ with TSP for internal referencing in the subsequent NMR analysis. For quality control, all fractions were analyzed with the dual receptor affinity assay.

2.3 On-line liquid chromatography – high resolution mass spectrometry dual receptor affinity detection system

A Shimadzu LCMS-IT-TOF in combination with a SPD20M photodiode array (PDA) detector was used for the liquid chromatography-high resolution mass spectrometry part of the screening assay while four additional Shimadzu LC10ADvp pumps and two Shimadzu RF10-XL fluorescence detectors were applied for the homogeneous dual receptor affinity assessment (Figure 1). The homogeneous dual receptor affinity detection system was developed based on the single line assay described by van Liempd *et al.* [11]. In brief, the receptor solution is added post column to the flow enabling the binding potential ligands in a 25 µL reaction coil made from PEEK, thermo stated at 37°C. Subsequently, the fluorescent tracer (coumestrol) is added to bind all free binding sites of the receptor and thereby enabling the detection of fluorescence enhancement of the receptor-tracer complex.

Samples were injected using a Shimadzu SIL20 auto-injector. Separation was performed on a Waters Xterra C18MS column (2.1 x 100 mm, 3.5 µm particles) equipped with a guard column. Mobile phase A2 consisted of 99% H₂O and 1% methanol, mobile phase B2 consisted of 1 % H₂O and 99 % methanol. The following gradient was applied with a flow-rate of 125 µL/min: 0-3 minutes isocratic at 35 % B2, 3-45 minutes linear increase to 100 % B2, where kept constant for 5 minutes and followed by a quick descent in 30 seconds to 35 % B2 for 10 minutes of equilibration prior to the next analysis. After separation, the

flow was directed into both biochemical assays and the IT-TOF mass spectrometer by a flow-split made from fused silica (30 μm I.D. 375 μm O.D. from BGB Analytics AG, Boeckten, Switzerland) introducing 100 $\mu\text{L}/\text{min}$ into the photodiode array and mass spectrometer and directing 12.5 $\mu\text{L}/\text{min}$ of the flow into each of the lines for the homogeneous receptor affinity detection system (Figure 1). Receptor solutions of 15 nM were infused with 100 $\mu\text{L}/\text{min}$ via 100 mL superloops (made in house by the VU University - Fine Mechanical Workshop). The receptor solution and column effluent were incubated at 37°C in a 25 μL reaction coil made from PEEK. Subsequently, coumestrol solutions of 433 nM for hER α and 144 nM for hER β were infused with 100 $\mu\text{L}/\text{min}$ via two other superloops and incubated at 37°C in a 20 μL reaction coil made from PEEK before being directed to the fluorescence detectors ($\lambda_{\text{ex}} = 340 \text{ nm}$, $\lambda_{\text{em}} = 410 \text{ nm}$). Receptor and tracer solutions were in 10 mM phosphate buffer pH 7.4 containing 150 mM NaCl and 0.4 mg/mL ELISA Blocking Reagent to prevent non-specific binding. All superloops containing receptor and tracer solutions were kept on ice.

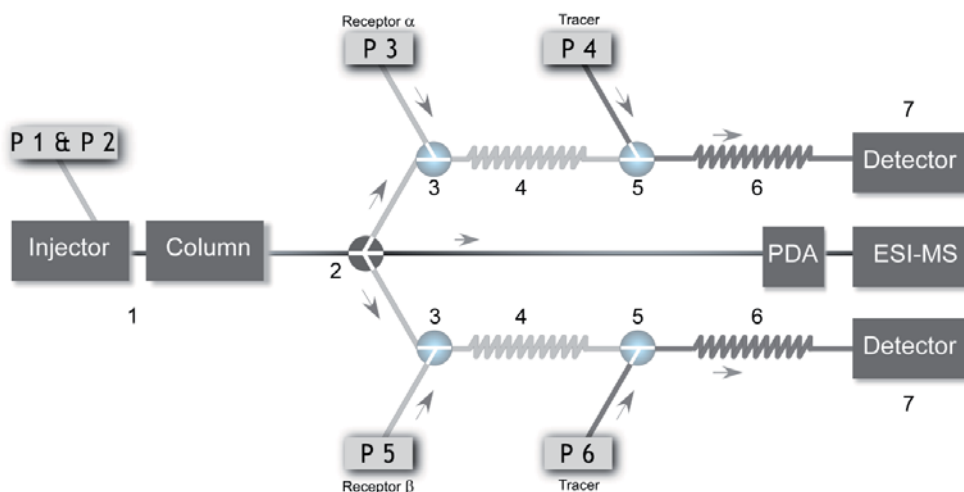


Figure 1: Configuration of the on-line liquid chromatography – high resolution mass spectrometry dual receptor affinity detection (RAD) system: (1) Sample injection followed by gradient HPLC [P1& P2]; (2) Splitting the flow to PDA-ESI-MS and the two RADs (hER α and hER β); (3) Infusing hER α and hER β via superloops and reagent pumps [P3&P5] (4) Binding of receptor and potential ligands in reaction coils; (5) Infusion of tracer solution via superloops and reagent pumps [P4&P6]; (6) Binding of fluorescent tracer to receptor in reaction coils; (7) Detection of receptor-tracer complex by fluorescence.

2.4 Mass spectrometer settings

A Shimadzu IT-TOF instrument equipped with an electrospray (ESI) source was operated in the data dependent acquisition mode switching between full-spectrum MS and MS-MS modes and alternating between positive and negative ion mode (100 ms cycles). The temperature of the heating block and curved desolvation line was set to 200 °C. Interface voltage was set at 5 kV for positive ion mode and - 4.5 kV for negative, the nebulizing gas flow (nitrogen, 99.999 % purity, Praxair, Oevel, Belgium) was 1.5 L/min and drying gas was applied at a rate of 10 L/min. Full MS scans were acquired between m/z 125 and 650. MS-MS data was obtained automatically with 10 ms accumulation time between m/z 100 and 650 and a precursor isolation width of 1.0 using Argon (99.9995 % purity, Praxair) as collision gas.

2.5 Validation of the screening methodology

The time correlation between assay response and response in the mass spectrometer was optimized with flow injection analysis. The length of the peek tubing between PDA and MS was adjusted in such a way, that it enabled direct correlation between the signals in MS and the biochemical detection system. To verify the time alignment, an incubation sample of estriol with cytochrome P450 BM3 M11_{his} was injected into the complete system. The dual receptor assay was performed in parallel using two times hER α to verify equal responses in both assay lines. After this, hER β was implemented in the screening assay. A dilution range of the natural ligand 17 β -E2 was made and injected in a 10 minutes LC run, isocratic elution at 70 % B2. As a quality control sample within every sequence of ten runs, 5 μ l of 17 β -E2 were injected at a concentration of 50 μ M. This also enabled the determination of intra and inter-day variability of the assay response.

2.6 NMR measurements

NMR analysis was performed on a Bruker DRX 600 MHz spectrometer equipped with a 5-mm TXI cryo probe and a Varian Inova 500MHz spectrometer equipped with a 5-mm probe both operating at 298K. The three-dimensional structure determination of the products is based on ^1H - ^1H -DQF-COSY, ^1H - ^{13}C -HSQC, ^1H - ^{13}C -HMQC, NOESY and/or 1D-proton spectra. The proton and carbon chemical shifts were referenced to the internal reference TSP (proton, δ =0.00ppm; carbon,

$\delta=0.00\text{ppm}$). All chemical entities extracted from the preparative scale biosynthesis were first analyzed and afterwards quantified by addition of a known concentration maleic acid (20 μg) [34]. Data were processed using ACDLabs software, version 11.03.

3. Results and Discussion

3.1 Validation of the screening methodology

The implementation of target-ligand interactions in an analytical methodology requires a different approach in validation from the approach taken for conventional bioanalytical methods. The use of hyphenated screening assays in the hit to lead selection process has to provide answers on the novelty of a compound and whether it is active or inactive against the used drug-target. Validation of the approach is necessary to assess the characteristics of the products formed. Accurate time alignment between the biological assay and the identification technique used, in this case the mass spectrometer is the most crucial factor. As an example, Figure 2 shows the alignment of the biochemical-detection assay and the extracted ion traces obtained in parallel. Clearly, the correlation between extracted ion chromatograms and the bioassay signal enables the positive identification of the m/z associated with the compounds binding to the biological target, not only in flow injection mode but more importantly, in a gradient elution setting. The correlation of traces in the bioassay (peak minima) and MS (peak maxima) were always within 6 seconds.

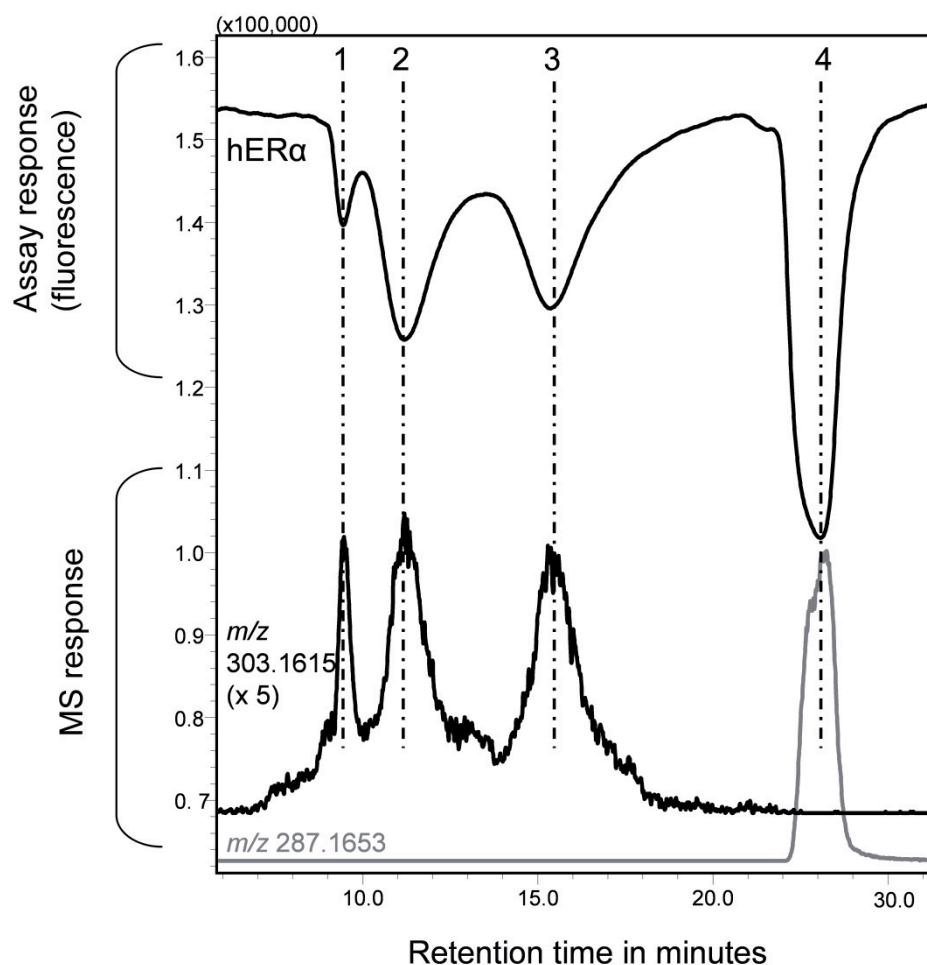


Figure 2: Correlation between biochemical detection (hER α , upper trace) and extracted ion traces from the mass-spectrometric detection (black m/z 303.1615 x 5 and gray m/z 287.1653, base line shifted) in an incubation mixture using estriol as substrate for cytochrome P450 BM3 M11_{his} biocatalysis.

The biological response in both assay lines was validated by using two times hER α subtype for both receptor affinity detection systems. With the injection of different concentrations (10 μ L of 1 nM-100 μ M) of the natural occurring ligand 17 β -E2 into the complete system, the responses recorded for both assay lines were similar and thus proving the validity of screening selectivity for both

receptors. Subsequently, hER α and hER β were mutually implemented in the screening assay and 17 β -E2 was injected to verify the responses for screening selectivity in receptor affinity (Figure 3). Inter and intra day variation of the signals in the dual RAD system was less than 5 % and the detection limit for 17 β -E2 was 4.7 ± 0.5 nM, which is well within limits reported for other hyphenated screening methods [10,35]. With the MS settings used for the screening of biosynthesis products, the detection limit for 17 β -E2 was 40 nM in negative ESI mode. This rather high detection limit compared to the bioassay is caused by the fact that 17 β -E2 is a very high affinity ligand for the hER subtypes and that this compound belongs to a steroid class that is reported to be ionized very inefficiently [36] in a non-derivatized form

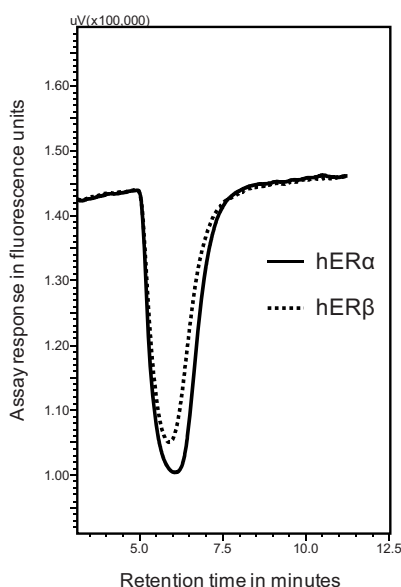


Figure 3: hER α (solid) and hER β (dashed) affinity responses of the on-line dual receptor affinity detection system after injection of 5 μ L 17 β -E2 (50 μ M).

3.2 Metabolite identification

3.2.1 Analytical scale biosynthesis and analysis

More than 85 products were formed using a test set of six compounds containing naturally occurring estrogens (17 α -E2 and 17 β -E2), synthetic estrogen receptor modulators and prodrugs (EE2, NET) and synthetic compounds from the Schering-Plough library (Org X and Org Y). Of these 85 products, 55 compounds were

identified as hits towards one or both hER's. Mainly, active species are (di)hydroxylated forms of the used scaffold while other active species are dehydrogenated or a combination of both. Figure 4 shows the typical affinity profiles obtained with our screening platform for the NET related structures, profiles obtained for EE2, 17 β -E2 and Org X are depicted in Figure S2. All products were characterized based on HR-MS data and affinity towards both receptors.

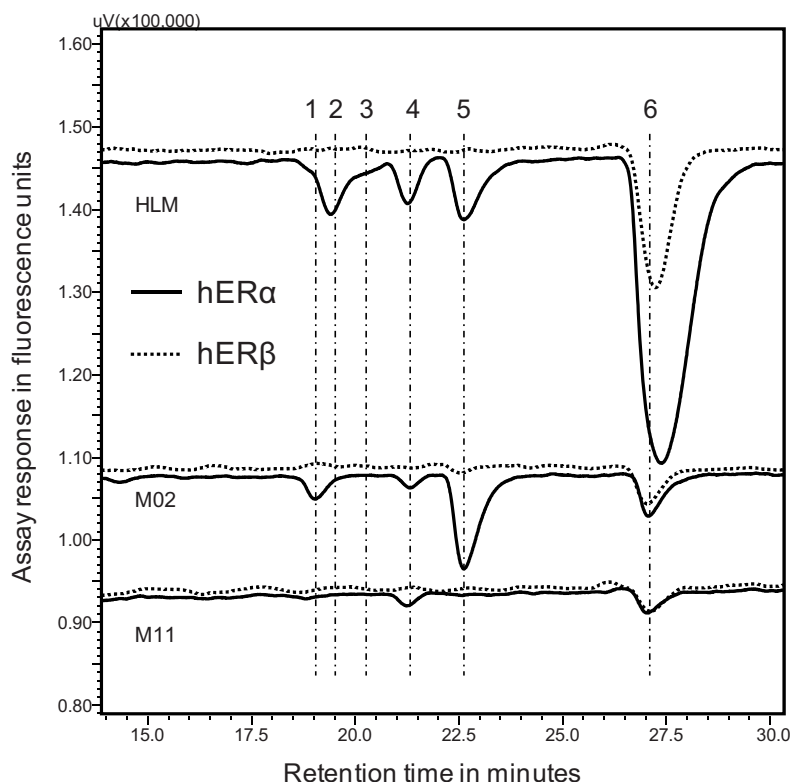


Figure 4: hER α (solid) and hER β (dashed) affinity traces obtained with the on-line LC-MS dual RAD system of reaction mixtures produced by HLM, M02_{his} and M11_{his} mutants using NET as scaffold. (1) M02-6; (2) metabolite with m/z 311.1664 ([M-H]⁻); (3) metabolite with m/z 311.1658 ([M-H]⁻); (4) M02-7; (5) M02-8; (6) M02-9. Detailed information on the compounds can be found in Table 1. Structures of all identified compounds are depicted in Figure 5.

Table 1 provides an inventory of all formed NET related products and their characteristics. MS detection enabled the proposal of the chemical formula, in

positive ion mode based on $[M+H]^+$, $[M+Na]^+$ and / or $[M+K]^+$ and for some compounds in the negative ion mode based on $[M-H]^-$. The PDA enables the determination of UV absorption maxima on all products formed. Based on the assumption that products with non-altered absorption profiles still have the same molar extinction coefficient, the quantity of new products can be estimated based on calibration curves obtained from available standards of the scaffold molecules. Relative affinity can therefore be rated by relating UV absorption of the produced compounds with their response in the biochemical assays, providing a quick evaluation of affinity towards the drug target(s) to support the medicinal chemistry follow up in lead optimization. A table with all products formed using 17 α -E2, 17 β -E2, EE2, Org X and Org Y as scaffold molecules can be found in the supplementary information (Table S6).

Table 1: Characterization of products using NET as scaffold. *m/z* indicated in bold represent the most abundant ion used for identification. Relative receptor affinity was ranked by relating UV absorbance and biochemical assay responses; - = no affinity, + = low affinity, ++ = moderate affinity and +++ = high affinity.

Product Identification NET analogues					Enzymatic System			Receptor Binding	
		High Resolution - MS							
<i>T_r</i> (min)	Modification	<i>m/z</i> positive ion mode	<i>m/z</i> negative ion mode	Code	M11 _{his}	M02 _{his}	HLM	hER α	hER β
8.8	+ 2 x O	[M+H]⁺ = 331.1914	[M-H] ⁻ = 329.1770	M02-1	-	✓	-	-	-
10.0	+ 2 x O		[M-H] ⁻ = 329.1762		-	-	✓	-	-
11.2	+ 2 x O	[M+H]⁺ = 331.1912		M02-2	-	✓	-	-	-
13.8	+ 2 x O	[M+Na]⁺ = 353.1713	[M-H] ⁻ = 329.1771	M02-3	-	✓	✓	-	-
16.0	+ O	[M+H]⁺ = 315.1968	[M-H] ⁻ = 313.1817	M02-4	✓	✓	✓	-	-
17.1	+ O	[M+H]⁺ = 315.1964		M02-5	✓	✓	✓	-	-
18.5	- 2H + O	[M+H]⁺ = 313.1810	[M-H] ⁻ = 311.1652		-	-	✓	-	-
18.9	+ O	[M+H] ⁺ = 315.1961	[M-H]⁻ = 313.1816	M02-6	-	✓	✓	+	-
19.1	- 2H + O		[M-H] ⁻ = 311.1664		-	-	✓	++	-
19.9	- 2H + O		[M-H] ⁻ = 311.1658		-	-	✓	++	-
20.9	+ O	[M+H]⁺ = 315.1965	[M-H] ⁻ = 313.1818	M02-7	✓	✓	✓	+	-
22.0	+ O	[M+H]⁺ = 315.1962	[M-H] ⁻ = 313.1819		-	-	✓	-	-
22.6	- 2H + O		[M-H] ⁻ = 311.1652	M02-8	-	✓	✓	+++	-
25.4	- 2H	[M+H] ⁺ = 297.1852			-	-	✓	-	+
26.1	NET	[M+H]⁺ = 299.2019	[M-H] ⁻ = 297.1872		✓	✓	✓	-	-
27.0	- 2H		[M-H] ⁻ = 295.1714	M02-9	✓	✓	✓	+++	+++

As recently described by Pozo *et al.* [37], the collision induced dissociation of 3-keto steroids after ESI results in complex product ion spectra. Characteristic fragment ions, indicating the site of modification of the steroid structure, would significantly facilitate metabolite identification. For example, in the case of the hydroxy metabolites M02-4 and M02-7 of norethisterone, elaborate spectra interpretation and comparison to the spectrum of the parent drug revealed that the m/z of several fragments containing the steroid D ring were incremented with the mass of an oxygen atom or reduced by the mass of two protons (due to water loss prior to the formation of that fragment) from the fragments of NET. Upon hydroxylation, an additional cleavage of H₂O is observed in the D ring, indicating the site of biotransformation and thereby significantly accelerating the NMR identification.

3.2.2 Preparative scale biosynthesis and NMR analysis

Full structure characterization was done on the products obtained from a preparative scale biosynthesis with NET as a substrate for biocatalysis. Using the highly active M02_{his} mutant, a set of compounds was generated for isolation and identification. All fractions obtained from the preparative LC analysis were screened for purity and affinity in the dual receptor affinity detection system and were simultaneously measured with NMR for full structure characterization. Figure 5 depicts the structures of the compounds generated by the large scale biosynthesis utilizing the P450 BM3 M02_{his} mutant.

The main product is the 16 β -hydroxylated product (M02-7), having a fairly low affinity towards hER α and no affinity towards hER β . The second most abundant product is 15 β -hydroxy NET (M02-4), which has no affinity for either of the receptors. The position of the hydroxyl group in the isolated biosynthesis products was confirmed and determined by comparing the ¹³C-¹H correlation spectra of the products with the spectra of NET. The analysis of M02-7 is described in the supplementary information as an example. NOESY spectra, as well as J-coupling patterns in the 1D ¹H spectra [38], were used to determine the 3D structures of the newly formed chemical entities. Interestingly, hydroxylation always occurred at the β position either at the A ring or D ring of the steroid, not at the α position or at the B or C ring. Jacobsen *et al.* found similar results when

they studied the hydroxylation of testosterone, progesterone and androstenedione by CYP6A1 [39].

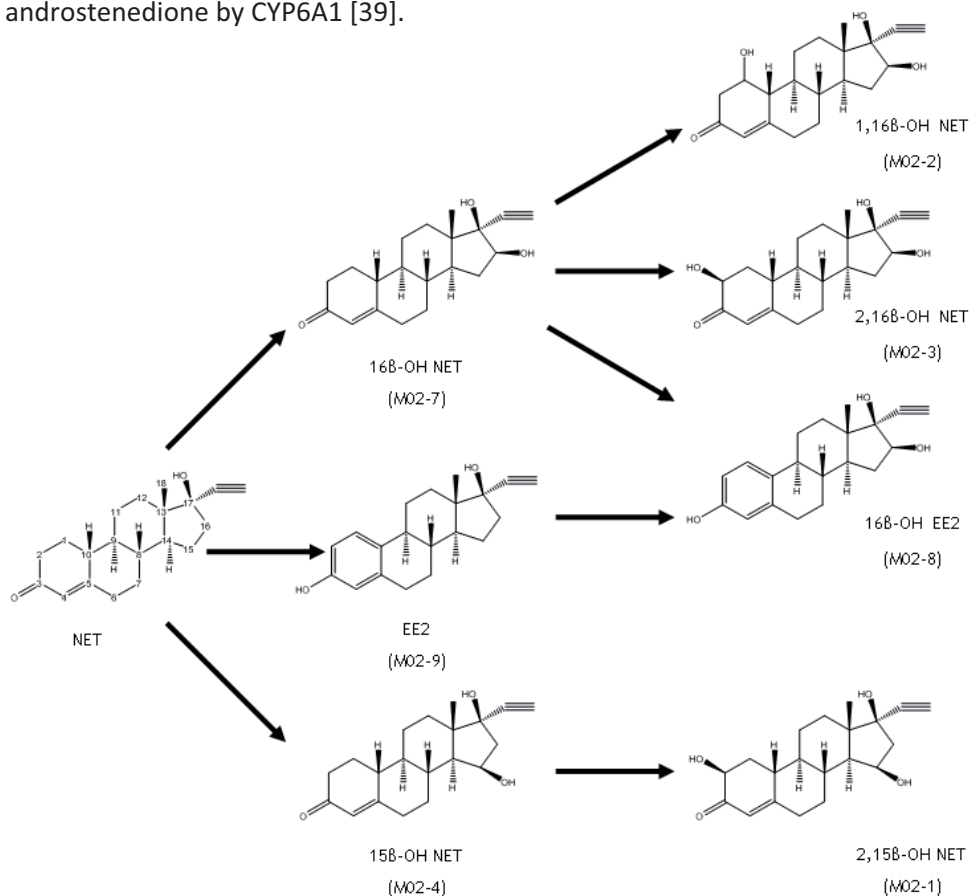


Figure 5: Products derived from biosynthesis with cytochrome P450 BM3 M02_{his} using norethisterone as scaffold.

The hydroxylation at position 1 in M02-2 is assumed to be at the β position but this could not be confirmed unambiguously by NMR due to the combination of low concentration of the compound and impurities present in the sample. NOESY spectra indicated that compounds with a hydroxyl at position 2, M02-1 and M02-3, have a folded A ring, identified by a strong NOE interaction between the proton at position 2 and the proton at position 9. This folding of the A ring into an inverted half chair conformation has been shown to be a common occurrence in steroids with a hydroxyl at this position [40-42]. We were not able to determine the structures of M02-5 and M02-6. M02-5 is an inactive compound and the

amount produced is insufficient for NMR experiments, HR-MS data indicates that the scaffold is hydroxylated. Product M02-6 degraded during the NMR data acquisition. Hence, based on the HR-MS, UV absorption data and its instability, this compound is probably the instable 4,5-epoxy-NET [43].

Interestingly, the M02_{his} mutant was able to form most of the active human relevant metabolites, albeit in different ratios. Recently, Korhonen *et al.* reported the elucidation of human CYPs involved in the metabolism of NET and proposed identities based on high resolution MS in positive ion mode and fragmentation studies with a triple quadrupole MS [44]. Our findings are in accordance with the published results, but due to the additional dimensions of our receptor affinity detection system and the utilization of fast polarity switching in full MS mode, we were able to identify additional metabolites with biological relevance. Compound M02-9 at *t_r* 27.0 min has high affinity for both Estrogen Receptor subtypes and is difficult to detect in positive ion mode. Although present in low concentrations, this entity is readily detected in negative ion mode and has a $[M-H]^-$ with *m/z* 295.1714 derived from dehydrogenation of NET or the loss of H₂O from a hydroxy-product. Based on retention time, affinity to both receptor subtypes and HR-MS/MS spectral library matching, we identified this product as EE2. Although not frequently described, NET can be metabolized in humans to EE2 via hydroxylation of the A ring and a subsequent loss of H₂O introducing aromaticity in the A ring. This effect was first described by Yamamoto *et al.* [45,46], whereas Kuhl *et al.* recently discussed clinical implications of the formation of this potent estrogenic active metabolite in humans [47]. As described by Ma *et al.* [36], aromatization of the A ring of a steroid, and the inherent loss of the 3 keto,4-ene conjugation, changes the ESI significantly. Additionally, the UV absorption maximum shifts from 244 nm to 285 nm indicating the introduction of an aromatic moiety. These two properties provide an extra tool in structure characterization of steroids.

Compound M02-08 is produced by the human liver microsomes as well as by the mutant M02_{his} and has high affinity towards the hER α subtype. This compound is formed in amounts too low to be identified by NMR but is easily detected in negative ion mode and has a $[M-H]^-$ with *m/z* 311.1652 (HR-MS spectrum in Figure S3). Additional experiments were done to identify the structure of M02-08. Since M02-08 is a secondary reaction product, the main product M02-07 was used as a substrate in analytical scale incubations with the mutants M02_{his} and M11_{his}.

The data (Figure S4) confirms that M02-08 is formed from M02-07 via a similar mechanism as norethisterone is converted into 17 α -ethinyloestradiol, namely via hydroxylation on the C2 and subsequent loss of H₂O. Based on these findings as well as on the change of its ionization properties [36] and UV absorption profile (Figure S5), M02-08 can be identified as 16 β -OH EE2. Other compounds were found to be secondary products as well, mainly being β -hydroxylated products at the C1, C2, C15 and C16 positions.

2 β , 15 β -dihydroxy norethisterone (M02-1): 9 μ g, ESI: m/z 331.1914 [M+H]⁺, m/z 329.1770 [M-H]⁻, ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm: 0.95 - 1.04 (m, 2 H, 9 α , 7 α) 1.00 (s, 3 H, CH₃) 1.22 - 1.29 (m, 2 H, 11 β , 14 α) 1.49 - 1.55 (m, 2 H, 12 α , 12 β) 1.57 - 1.67 (m, 1 H, 8 β) 1.80 - 1.85 (m, 1 H, 11 α) 1.84 - 1.91 (m, 2 H, 1 α , 16 β) 1.96 - 2.04 (m, 1 H, 1 β) 2.17 - 2.24 (m, 1 H, 10 β) 2.28 - 2.34 (m, 1 H, 7 β) 2.38 - 2.46 (m, 2 H, 6 α , 6 β) 2.46 - 2.54 (m, 1 H, 16 α) 3.90 (dd, 1 H, 2 α) 3.99 - 4.05 (m, 1 H, 15 α) 5.67 (s, 1 H, 4).

1, 16 β -dihydroxy norethisterone (M02-2): 6 μ g, ESI : m/z 331.1912 [M+H]⁺, ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm: 0.81 (s, 3 H, CH₃) 0.83 - 0.91 (m, 1 H, 9 α) 0.96 - 1.04 (m, 1 H, 7 α) 1.04 - 1.13 (m, 1 H, 15 β) 1.12 - 1.19 (m, 1 H, 14 α) 1.30 - 1.39 (m, 1 H, 11 β) 1.39 - 1.47 (m, 1 H, 8 β) 1.48 - 1.54 (m, 1 H, 12 β) 1.59 - 1.69 (m, 1 H, 12 α) 1.82 - 1.93 (m, 2 H, 7 β , 11 α) 2.07 - 2.15 (m, 1 H, 15 α) 2.12 - 2.23 (m, 1 H, 10 β) 2.20 - 2.26 (m, 1 H, 6 β) 2.35 - 2.41 (m, 1 H, 6 α) 2.45 - 2.55 (m, 2 H, 2 α , 2 β) 3.95 (m, 1 H, 16 α) 4.21 (m, 1 H, 1) 4.65 (b, 1 H, OH17 β) 4.83 (b, 1 H, OH1) 5.28 (b, 1 H, OH16 β) 5.68 (s, 1 H, 4).

2 β , 16 β -dihydroxy norethisterone (M02-3): 44 μ g, ESI: m/z 353.1713 [M+Na]⁺, m/z 329.1771 [M-H]⁻, ¹H NMR (500 MHz, DMSO-*d*₆) δ ppm: 0.82: (s, 3 H, methyl) 0.93 - 1.05 (m, 2 H, 7 α , 9 α) 1.08 - 1.16 (m, 1 H, 15 β) 1.16 - 1.24 (m, 1 H, 14 α) 1.23 - 1.33 (m, 1 H, 11 β) 1.35 - 1.45 (m, 1 H, 8 β) 1.50 - 1.57 (m, 1 H, 12 β) 1.62 - 1.71 (m, 1 H, 12 α) 1.80 - 1.90 (m, 3 H, 1 α , 7 β , 11 α) 1.95 - 2.03 (m, 1 H, 1 β) 2.09 - 2.16 (m, 1 H, 15 α) 2.21 - 2.28 (m, 1 H, 6 β) 2.28 - 2.32 (m, 1 H, 10 β) 2.38 - 2.44 (m, 1 H, 6 α) 3.87 - 3.93 (m, 1 H, 2 α) 3.94 - 3.99 (m, 1 H, 16 α) 4.55 - 4.74 (b, 1 H, OH) 5.11 - 5.41 (br, 2 H, 2 β OH, 16 β OH) 5.67 (s, 1 H, 4). ¹³C NMR (500 MHz, DMSO-*d*₆) δ ppm: 12.6 (C18) 25.6 (C11) 31.8 (C7) 32.7 (C1) 33.0 (C12) 34.1 (C15) 34.8 (C6) 40.21 (C8) 40.4 (C10) 45.0 (C14) 47.8 (C9) 68.2 (C2) 76.2 (C16) 120.3 (C4).

15 β -hydroxy norethisterone (M02-4): 393 μ g, ESI: m/z 315.1968 $[M+H]^+$, m/z 313.1817

$[M-H]^-$, 1H NMR (500 MHz, DMSO- d_6) δ ppm: 0.73 - 0.85 (m, 1 H, 9 α) 0.90 - 0.98 (m, 1 H, 7 α) 1.01 (s, 3 H, methyl) 1.16 - 1.21 (m, 1 H, 11 β) 1.22 - 1.28 (m, 1 H, 14 α) 1.41 - 1.47 (m, 1 H, 1 α) 1.48 - 1.54 (m, 1 H, 12 β) 1.54 - 1.60 (m, 1 H, 12 α) 1.61 - 1.70 (m, 1 H, 8 β) 1.76 - 1.84 (m, 1 H, 11 α) 1.84 - 1.91 (m, 1 H, 16 β) 2.07 - 2.16 (m, 1 H, 7 β) 2.17 - 2.26 (m, 4 H, 1 β , 2 β , 2 α , 10) 2.26 - 2.33 (m, 1 H, 6 β) 2.39 - 2.47 (m, 1 H, 6 α) 2.53 - 2.59 (m, 1 H, 16 α) 4.00 - 4.10 (m, 1 H, 15 α) 4.45 - 4.57 (d, 1 H, 15 β OH) 5.29 - 5.41 (s, 1 H, OH) 5.69 - 5.78 (s, 1 H, 4). ^{13}C NMR (600 MHz, DMSO- d_6) δ ppm: 15.0 (C18), 25.5 (C11), 26.0 (C1), 29.7 (C7), 33.4 (C12), 34.7 (C6), 36.0 (C2), 36.1 (C8), 42.0 (C10), 49.1 (C9), 51.6 (C16), 53.5 (C14), 66.5 (C15), 123.8 (C4). Previously identified with NMR by Ambrus *et al.* [48].

16 β -hydroxy norethisterone (M02-7): 485 μ g, ESI: m/z 315.1965 $[M+H]^+$, m/z 313.1818 $[M-H]^-$, 1H NMR (500 MHz, DMSO- d_6) δ ppm: 0.72 - 0.85 (m, 1 H, 9 α) 0.81 (s, 3 H, methyl) 0.92 - 1.02 (m, 1 H, 7 α) 1.08 - 1.15 (m, 1 H, 15 β) 1.15 - 1.20 (m, 1 H, 14 α) 1.20 - 1.27 (m, 1 H, 11 β) 1.34 - 1.42 (m, 1 H, 8 β) 1.43 - 1.49 (m, 1 H, 1 α) 1.50 - 1.55 (m, 1 H, 12 β) 1.59 - 1.67 (m, 1 H, 12 α) 1.74 - 1.79 (m, 1 H, 7 β) 1.79 - 1.84 (m, 1 H, 11 α) 2.12 - 2.18 (m, 1 H, 15 α) 2.15 - 2.18 (m, 1 H, 10 β) 2.17 - 2.21 (m, 1 H, 1 β) 2.20 - 2.24 (m, 2 H, 2 α , 2 β) 2.23 - 2.32 (m, 1 H, 6 β) 2.36 - 2.45 (m, 1 H, 6 α) 3.94 - 4.00 (m, 1 H, 16 α) 4.64 (s, 1 H, OH) 5.26 - 5.27 (d, 1 H, 16 β OH) 5.72 (s, 1 H, 4). ^{13}C NMR (600 MHz, DMSO- d_6) δ ppm: 12.6 (C18), 25.2 (C11), 25.7 (C1), 30.3 (C7), 33.4 (C12), 34.3 (C15), 34.5 (C6), 35.8 (C2), 39.7 (C8), 41.6 (C10), 45.2 (C14), 48.8 (C9), 76.3 (C16), 124.0 (C4).

Norethisterone (NET): ESI: m/z 299.2019 $[M+H]^+$, m/z 297.1872 $[M-H]^-$, 1H NMR (600 MHz, DMSO- d_6) δ ppm: 0.72 - 0.78 (m, 1 H, 9 α) 0.79 (s, 3 H, methyl) 0.92 - 1.01 (m, 1 H, 7 α) 1.17 - 1.25 (m, 1 H, 11 β) 1.25 - 1.31 (m, 1 H, 15 β) 1.31 - 1.38 (m, 1 H, 8 β) 1.38 - 1.45 (m, 1 H, 14 α) 1.45 - 1.52 (m, 1 H, 1 α) 1.53 - 1.59 (m, 1 H, 15 α) 1.58 - 1.65 (m, 2 H, 12 α , 12 β) 1.73 - 1.79 (m, 1 H, 7 β) 1.80 - 1.84 (m, 1 H, 11 α) 1.83 - 1.88 (m, 1 H, 16 β) 2.02 - 2.18 (m, 1 H, 16 α) 2.12 - 2.18 (m, 1 H, 10 β) 2.17 - 2.21 (m, 1 H, 1 β) 2.20 - 2.24 (m, 2 H, 2 α , 2 β) 2.23 - 2.31 (m, 1 H, 6 β) 2.39 - 2.45 (m, 1 H, 6 α) 5.33 (s, 1 H, OH) 5.69 - 5.78 (m, 1 H, 4). ^{13}C NMR (600 MHz, DMSO- d_6) δ ppm: 12.3 (C18), 22.5 (C15), 25.6 (C11), 26.0 (C1), 30.4 (C7), 32.3 (C12), 34.6 (C6), 36.1 (C2), 38.7 (C16), 40.3 (C8), 41.8 (C10), 48.8 (C14), 49.0 (C9), 124.0 (C4).

The NMR assignment of NET in chloroform- d_1 is previously reported by Sedee *et al.* [49].

3.3 Identification Approach

The approach followed in this work is based on two major parts, being the analytical scale biosynthesis and analysis utilizing high resolution MS(/MS) (I) and preparative biosynthesis followed by isolation and NMR (II). In general, the production of novel metabolites is assessed using analytical scale biosynthesis. This first part generates information on chemical structure and bioaffinity by utilizing the on-line LC–MS dual receptor affinity detection system. Specific profiles are obtained of all metabolites formed and include identifying factors such as accurate mass, retention time, relative affinity information, UV-absorption profiles and MS/MS fingerprints. When profiles match to the library, the identification with HR-MS/MS is sufficient for structure identification. Due to the multidimensional data matrix obtained on all compounds, the upscaling for a complete structure characterization is only limited to the interesting novel compounds generated by specific enzymatic systems. For these novel compounds, preparative scale biosynthesis was performed for subsequent preparative LC fractionation to determine the structure conformation by NMR while the process is monitored with the on-line dual receptor affinity detection system. Insufficient product formation (ca. < 1 μ g) disables *de novo* structure elucidation by NMR due to the detection limits of the NMR employed (magnetic fields up to 600 MHz). By using, for example, micro-droplet NMR, NMR on a chip or NMR at higher-field (800 MHz equipped with cryo-cooled probes), detection limits may be reduced to ca. 250 ng [50,51]. We further note that in contrast to *de novo* characterization, NMR identification of pure stable known library compounds lowers these detection limits. However, in our experience, full purity may often not be achieved in practice and mixtures with vastly different amounts may occur. Moreover, in steroids, complex J-coupling multiplets (doublets up to doublets of doublets of doublets of doublets) tend to occur, which leads to higher complexity and lower peak height per proton as compared to singlet resonances. These aspects made the given detection limits (250 to 1000 ng) approximate but still a practical guideline and can thus be referred to as ‘limit of spectroscopic

identification' (LOSI) as discussed recently by Wilson *et al.* for complicated hyphenated systems [52].

4. Conclusions

In total, the analytical scale biosynthesis produced over 85 compounds from 6 different starting templates. As an example, the NET related biotransformation products were subjected to full structure characterization. Reaction products of NET formed by M02_{his} were identified based on HR-MS(/MS) data, drug-target affinity profiles, UV absorption profiles and spectral library matching of standards. Of all NET related products generated by the M02_{his} enzymatic system, we were able to fully identify ~70% based on the different strategies and to obtain final structure confirmation by NMR data. Based on the full structure characterization of the M02_{his} produced norethisterone analogues, the majority of all products formed by the other enzymatic systems (M11_{his} and HLM) were identified by LC-HR-MS(/MS) experiments. In conclusion, by combining hyphenated screening assays, HR-MS(/MS) for fast, first line structure identification and NMR spectroscopy for full 3D structure confirmation we were able to determine, identify and characterize chemical entities produced by biosynthetic enzymes. This multidimensional screening approach enabled the detection of low abundant biotransformation products which were not suitable for detection in either one of its single components. The implementation of this approach in early stage drug discovery expands the toolbox of the medicinal chemist for the generation and optimization of lead compounds in a quick and informative way. A combination of compound library generation, information on metabolic soft spots and drug target selectivity is obtained in parallel and thereby saving valuable time and manpower. To improve the versatility and throughput of our approach, our current research includes the expansion of our mutant enzyme library, data stream clustering methods and MS/NMR based drug target affinity assays.

Acknowledgements

This research was performed within the framework of project D2-102 of the Dutch Top Institute Pharma. The authors thank Jeroen Lastdrager of the LACDR division of Molecular Toxicology, Ruud Aspers and Rianne Willems of Schering-Plough

Research Institute Oss for their technical assistance and helpful discussions. Shimadzu Benelux is acknowledged for their technical support.

References

- [1] F.P. Guengerich, *Nat. Rev. Drug Discovery* 1 (2002) 359.
- [2] C.R. Otey, G. Bandara, J. Lalonde, K. Takahashi, F.H. Arnold, *Biotechnol. Bioeng.* 93 (2006) 494.
- [3] W.Y. Li, J.L. Josephs, G.L. Skiles, W.G. Humphreys, *Drug Metab. Dispos.* 36 (2008) 721.
- [4] M. Zmijewski, T.A. Gillespie, D.A. Jackson, D.F. Schmidt, P. Yi, P. Kulanthaivel, *Drug Metab. Dispos.* 34 (2006) 925.
- [5] A.W. Munro, D.G. Leys, K.J. McLean, K.R. Marshall, T.W.B. Ost, S. Daff, C.S. Miles, S.K. Chapman, D.A. Lysek, C.C. Moser, C.C. Page, P.L. Dutton, *Trends Biochem. Sci.* 27 (2002) 250.
- [6] M. Landwehr, L. Hochrein, C.R. Otey, A. Kasrayan, J.E. Backvall, F.H. Arnold, *J. Am. Chem. Soc.* 128 (2006) 6058.
- [7] B.M.A. Lussenburg, L.C. Babel, N.P.E. Vermeulen, J.N.M. Commandeur, *Anal. Biochem.* 341 (2005) 148.
- [8] B.M.A. van Vugt-Lussenburg, M.C. Damsten, D.M. Maasdijk, N.P.E. Vermeulen, J.N.M. Commandeur, *Biochem. Biophys. Res. Commun.* 346 (2006) 810.
- [9] B.M.A. van Vugt-Lussenburg, E. Stjernschantz, J. Lastdrager, C. Oostenbrink, N.P.E. Vermeulen, J.N.M. Commandeur, *J. Med. Chem.* 50 (2007) 455.
- [10] A.J. Oosterkamp, M.T.V. Herraiz, H. Irth, U.R. Tjaden, J. van der Greef, *Anal. Chem.* 68 (1996) 1201.
- [11] S.M. van Liempd, J. Kool, W.M.A. Niessen, D.E. van Elswijk, H. Irth, N.P.E. Vermeulen, *Drug Metab. Dispos.* 34 (2006) 1640.
- [12] J. Kool, A. van Marle, S. Hulscher, M. Selman, D.J. van Iperen, K. van Altena, M. Gillard, R.A. Bakker, H. Irth, R. Leurs, N.P.E. Vermeulen, *J. Biomol. Screen.* 12 (2007) 1074.
- [13] T. Schenk, J. Breel, P. Koevoets, S. van den Berg, A.C. Hogenboom, H. Irth, U.R. Tjaden, J. van der Greef, *J. Biomol. Screen.* 8 (2003) 421.
- [14] A.J. Oosterkamp, R. van der Hoeven, W. Glassgen, B. Konig, U.R. Tjaden, J. van der Greef, H. Irth, *J. Chromatogr. B* 715 (1998) 331.

-
- [15] J. Reinen, J. Kool, N.P.E. Vermeulen, *Anal. Bioanal. Chem.* 390 (2008) 1987.
 - [16] A.R. de Boer, T. Letzel, H. Lingeman, H. Irth, *Anal. Bioanal. Chem.* 381 (2005) 647.
 - [17] A.R. de Boer, T. Letzel, D.A. van Elswijk, H. Lingeman, W.M.A. Niessen, H. Irth, *Anal. Chem.* 76 (2004) 3155.
 - [18] A.R. de Boer, J.M. Alcaide-Hidalgo, J.G. Krabbe, J. Kolkman, C.N.V. Boas, W.M.A. Niessen, H. Lingeman, H. Irth, *Anal. Chem.* 77 (2005) 7894.
 - [19] A.R. de Boer, H. Lingeman, W.M.A. Niessen, H. Irth, *Trac-Trends in Analytical Chemistry* 26 (2007) 867.
 - [20] C.F. de Jong, R.J.E. Derks, B. Bruyneel, W. Niessen, H. Irth, *J. Chromatogr.* 1112 (2006) 303.
 - [21] A.P. Watt, R.J. Mortishire-Smith, U. Gerhard, S.R. Thomas, *Curr. Opin. Drug Discovery Dev.* 6 (2003) 57.
 - [22] R.F. Staack, G. Hopfgartner, *Anal. Bioanal. Chem.* 388 (2007) 1365.
 - [23] S.G. Ma, M.S. Zhu, *Chem.-Biol. Interact.* 179 (2009) 25.
 - [24] L. Ridder, M. Wagener, *Chemmedchem* 3 (2008) 821.
 - [25] M.S. Zhu, L. Ma, D.L. Zhang, K. Ray, W.P. Zhao, W.G. Humphreys, G. Skiles, M. Sanders, H.Y. Zhang, *Drug Metab. Dispos.* 34 (2006) 1722.
 - [26] P.J. Zhu, W. Tong, K. Alton, S. Chowdhury, *Anal. Chem.* 81 (2009) 5910.
 - [27] Y. Landry, J.P. Gies, *Fundam. Clin. Pharmacol.* 22 (2008) 1.
 - [28] G.G.J.M. Kuiper, P.J. Shughrue, I. Merchenthaler, J.A. Gustafsson, *Front. Neuroendocrinol.* 19 (1998) 253.
 - [29] G.G.J.M. Kuiper, G.J.C.M. van den Bemd, J.P.T.M. van Leeuwen, J. *Endocrinol. Invest.* 22 (1999) 594.
 - [30] G.P. Skliris, E. Leygue, P.H. Watson, L.C. Murphy, *J. Steroid Biochem. Mol. Biol.* 109 (2008) 1.
 - [31] J. Hartman, A. Ström, J.-Å. Gustafsson, *Steroids* 74 (2009) 635.
 - [32] M.C. Damsten, B.M.A. van Vugt-Lussenburg, T. Zeldenthuis, J.S.B. de Vlieger, J.N.M. Commandeur, N.P.E. Vermeulen, *Chemico-Biological Interactions* 171 (2008) 96.
 - [33] S. Eiler, M. Gangloff, S. Duclaud, D. Moras, M. Ruff, *Protein Expression Purif.* 22 (2001) 165.
 - [34] A.A. Salem, H.A. Mossa, B.N. Barsoum, *J. Pharm. Biomed. Anal.* 41 (2006) 654.

-
- [35] U. Schobel, M. Frenay, D.A. Van Elswijk, J.M. McAndrews, K.R. Long, L.M. Olson, S.C. Bobzin, H. Irth, J. Biomol. Screen. 6 (2001) 291.
 - [36] Y.-C. Ma, H.-Y. Kim, J. Am. Soc. Mass Spectrom. 8 (1997) 1010.
 - [37] O.J. Pozo, P. Van Eenoo, K. Deventer, S. Grimalt, J.V. Sancho, F. Hernandez, F.T. Delbeke, Rapid Commun. Mass Spectrom. 22 (2008) 4009.
 - [38] D.N. Kirk, H.C. Toms, C. Douglas, K.A. White, K.E. Smith, S. Latif, R.W.P. Hubbard, J. Chem. Soc., Perkin Trans. 2 (1990) 1567.
 - [39] N.E. Jacobsen, K.E. Kover, M.B. Murataliev, R. Feyereisen, F.A. Walker, Magn. Reson. Chem. 44 (2006) 467.
 - [40] D. Colombo, P. Ferraboschi, L. Legnani, P. Prestileo, L. Toma, J. Steroid Biochem. Mol. Biol. 103 (2007) 163.
 - [41] W.L. Duax, C. Eger, Pokrywie.S, Y. Osawa, J. Med. Chem. 14 (1971) 295.
 - [42] K. Kuriyama, E. Kondo, K. Tori, Tetrahedron Lett. (1963) 1485.
 - [43] C.E. Cook, M.C. Dickey, Christen.Hd, Drug Metab. Dispos. 2 (1974) 58.
 - [44] T. Korhonen, M. Turpeinen, A. Tolonen, K. Laine, O. Pelkonen, J. Steroid Biochem. Mol. Biol. 110 (2008) 56.
 - [45] T. Yamamoto, C. Sakai, J. Yamaki, K. Takamori, S. Yoshiji, J. Kitawaki, M. Fujii, J. Yasuda, H. Honjo, H. Okada, Endocrinol. Jpn. 31 (1984) 277.
 - [46] T. Yamamoto, S. Yoshiji, J. Yasuda, K. Shiroshita, J. Kitawaki, M. Fujii, M. Urabe, H. Honjo, H. Okada, Endocrinol. Jpn. 33 (1986) 527.
 - [47] H. Kuhl, I. Wiegratz, Climacteric 10 (2007) 344.
 - [48] G. Ambrus, E. Szarka, I. Barta, K. Albrecht, G. Horvath, Acta Microbiol. Acad. Sci. Hung. 22 (1975) 453.
 - [49] A.G.J. Sedee, G.M.J.B. Vanhenegouwen, W. Guijt, C.A.G. Haasnoot, J. Chem. Soc., Perkin Trans. 2 (1984) 1755.
 - [50] Y.Q. Lin, S. Schiavo, J. Orjala, P. Vouros, R. Kautz, Anal. Chem. 80 (2008) 8045.
 - [51] J. Bart, A.J. Kolkman, A.J. Oosthoek-de Vries, K. Koch, P.J. Nieuwland, H. Janssen, J.P.J.M. van Bentum, K.A.M. Ampt, F.P.J.T. Rutjes, S.S. Wijmenga, H. Gardeniers, A.P.M. Kentgens, J. Am. Chem. Soc. 131 (2009) 5014.
 - [52] I.D. Wilson, U.A.T. Brinkman, Trac-Trends in Analytical Chemistry 26 (2007) 847.

Supporting Information

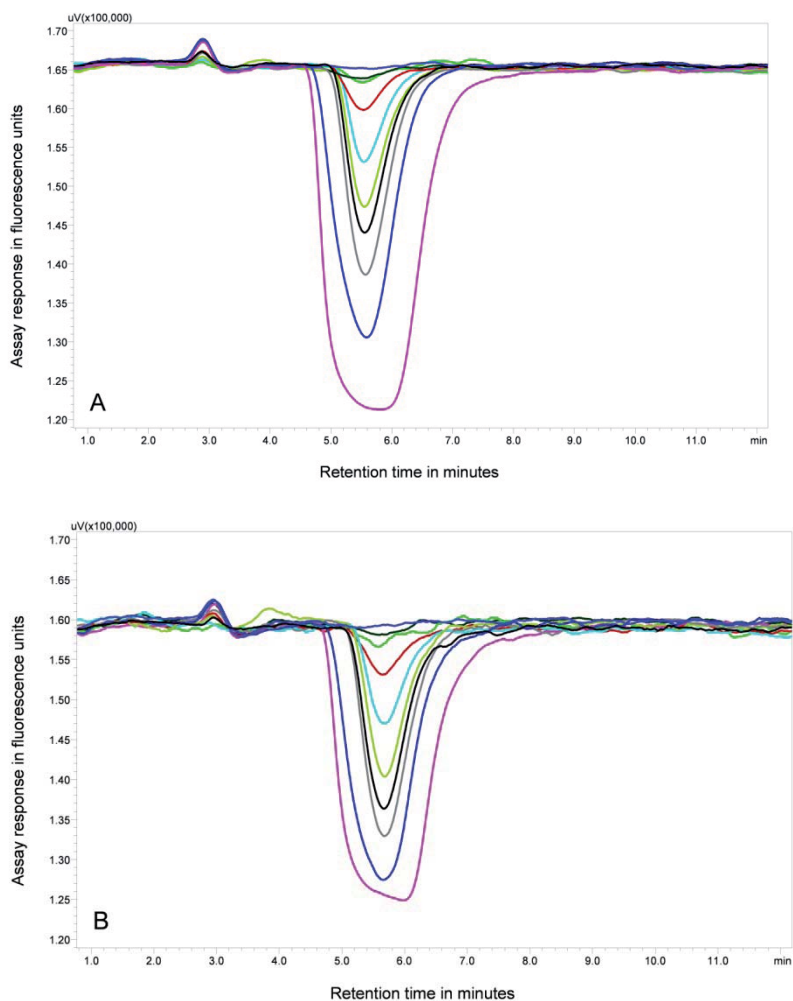


Figure S1: Biological response of the dual receptor affinity detection system, using both (A) hER α and (B) hER β with injections of 17 β -E2 (10 μ L of 1 nM-100 μ M).

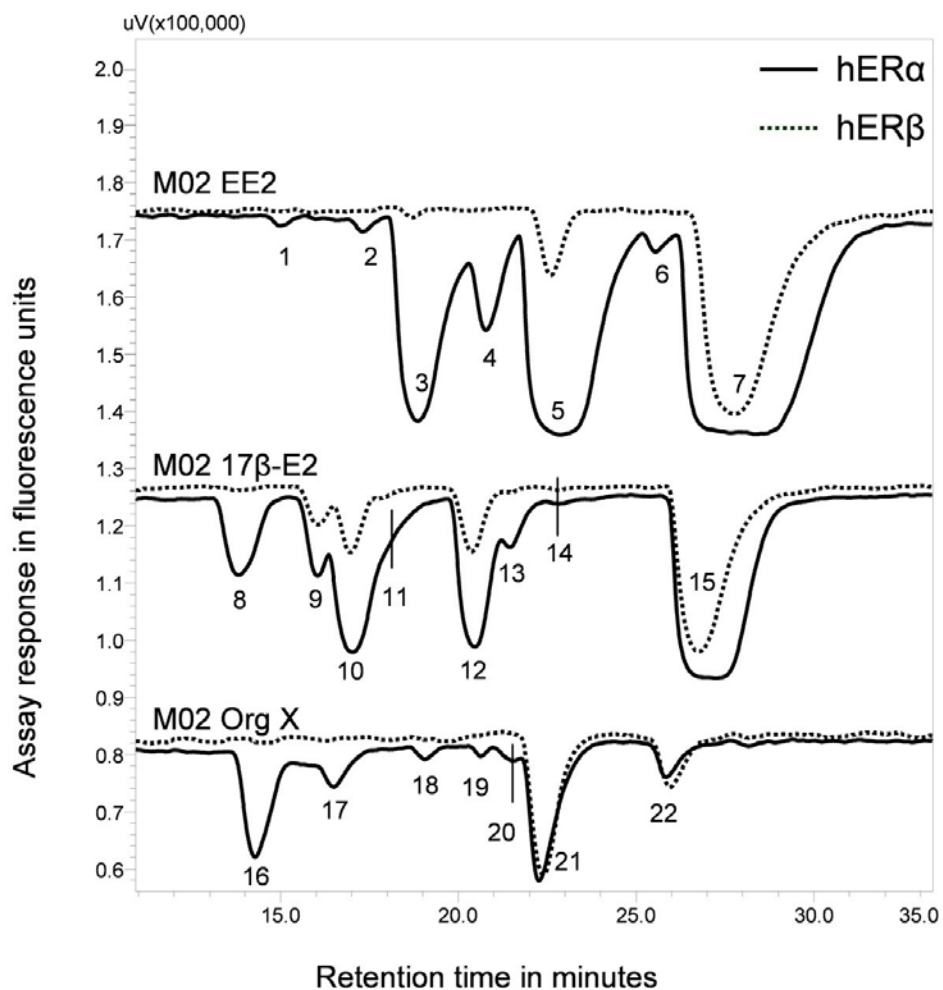


Figure S2: On-line LC-MS dual RAD hER α (solid) and hER β (dashed) affinity traces of reaction mixtures produced by M02_{his} using EE2, 17 β -E2 and Org X as scaffolds, (#) numbers indicate peaks as depicted in table S6.

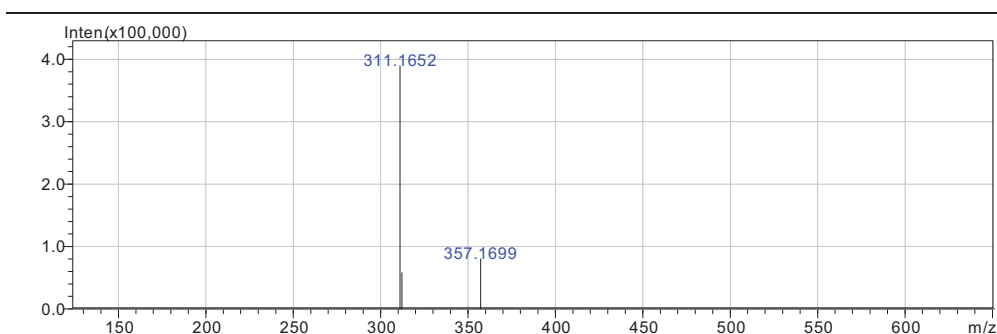


Figure S3: High resolution MS spectrum of 16β-OH EE2 (M02-8), in negative ion mode, showing $[M-H]^-$ (0.3 ppm) and $[M+HCOO]^-$ (2.2 ppm)

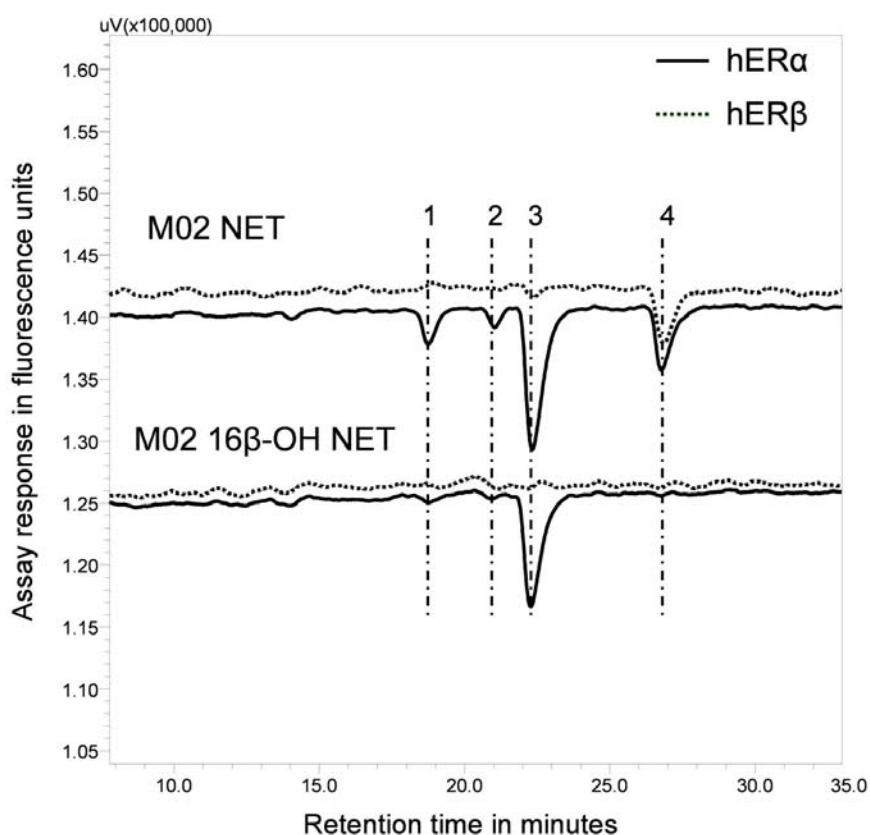


Figure S4: hERα (solid) and hERβ (dashed) affinity profiles of biosynthesis products by P450 BM3 M02_{his} using NET and of 16β-OH NET (M02-7) as scaffolds. (1) M02-6, (2) M02-7, (3) M02-8, (4) M02-9.

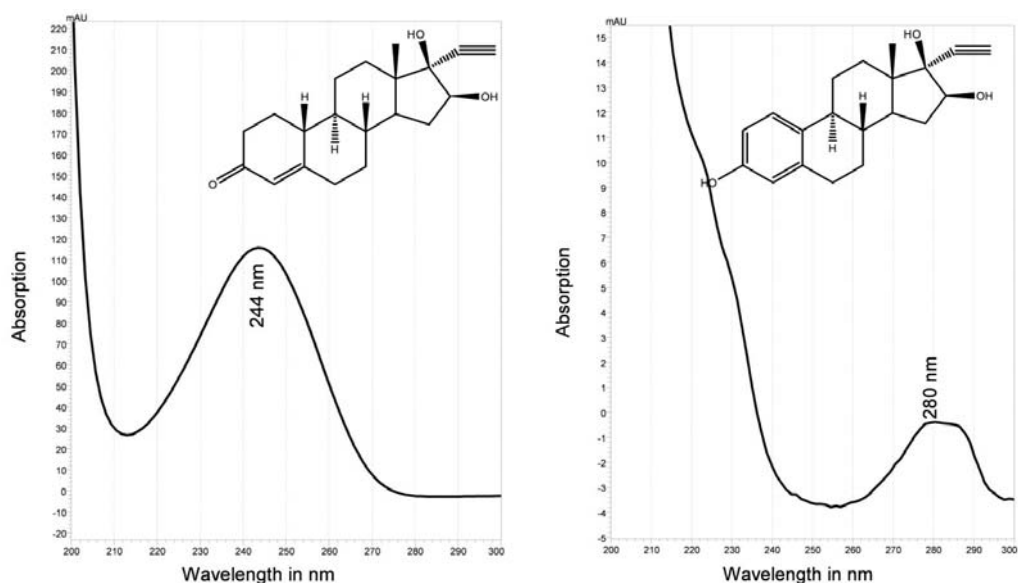


Figure S5: UV absorption profiles obtained in the screening of NET-products (left)

16β-OH NET (M02-7) and (right) 16β-OH EE2 (M02-8)

Table S6 (next page): Characterization of produced entities. (#) between brackets the peak number as indicated in figure S2; *m/z* values indicated in bold represent the most abundant ion used for identification, all *m/z* values mentioned for negative ion mode represent the $[M-H]^-$. Relative receptor affinity was ranked by relating UV-absorbance and biochemical-assay responses; - = no affinity, + = low affinity, ++ = moderate affinity and +++ = high affinity.

Determination and identification of estrogenic compounds

Product Identification				Enzymatic System			Receptor Binding	
T _r (min)	Modification	High Resolution - MS		M11 _{his}	M02 _{his}	HLM	hERα	hERβ
		m/z + ion mode	m/z – ion mode					
EE2								
11.2	+ O		311.1660	✓	✓	✓	–	–
14.1	+ O		311.1664	–	–	✓	++	–
(1) 15.0	+ O		311.1653	✓	✓	✓	++	–
16.5	+ O	[M+K] ⁺ = 351.1362	311.1658	–	–	✓	+	–
(2) 17.3	+ O	[M+K] ⁺ = 351.1371	311.1660	✓	✓	✓	+	–
(3) 18.7	+ O	[M+K] ⁺ = 351.1373	311.1666	✓	✓	✓	++	+
19.5	+ O	[M+K] ⁺ = 351.1367	311.1656	–	–	✓	+	–
(4) 20.7	- 2H + O	[M+H] ⁺ = 311.1632	309.1499	✓	✓	✓	+	–
21.5	+ O		311.1663	–	–	✓	++	–
(5) 22.2	+ O	[M+Na] ⁺ = 335.1632	311.1656	✓	✓	✓	++	+
22.9	- 2H + O		309.1503	✓	–	✓	+	+
24.0	+ 2H + O	[M+K] ⁺ = 353.1525	313.1813	✓	✓	–	–	–
24.2	+ O		311.1661	✓	–	✓	++	+
(6) 25.5	- 2H		293.1544	✓	✓	✓	++	+
(7) 26.8	Substrate		295.1709	✓	✓	✓	+++	+++
17β-E2								
6.2	+ 2 x O		303.1605	✓	✓	✓	–	–
8.3	- 2H + 2 x O	[M+K] ⁺ = 341.1149	301.1445	–	✓	–	–	–
9.6	- 2H + 2 x O	[M+K] ⁺ = 341.1150	301.1439	–	✓	–	–	–
9.9	+ 2 x O	[M+K] ⁺ = 343.1297	303.1594	✓	✓	–	–	–
10.6	+ 2 x O	[M+K] ⁺ = 343.1304	303.1607	✓	✓	–	–	–
11.6	- 2H + 2 x O		301.1445	–	✓	–	–	–
12.2	+ 2 x O		303.1607	–	–	✓	++	–
(8) 13.8	+ O		287.1652	✓	✓	–	++	–
15.0	+ O		287.1662	–	–	✓	–	–
(9) 15.9	+ O		287.1652	✓	✓	✓	++	++
(10) 16.8	+ O		287.1664	✓	✓	✓	+++	+++

T _r (min)	Modification	m/z + ion mode	m/z – ion mode	M11 _{his}	M02 _{his}	HLM	hERα	hERβ
(11) 17.8	- 2H + O	[M+K] ⁺ = 325.1195	285.1492	✓	✓	✓	++	–
18.8	- 2H + O		285.1505	–	–	✓	++	–
(12) 20.2	+ O		287.1651	✓	✓	✓	+++	+++
(13) 21.5	+ O		287.1658	✓	✓	✓	++	–
(14) 22.6	+ O		287.1663	✓	✓	✓	+++	+++
23.3	+ O		287.1662	–	–	✓	+++	+++
25.1	- 2H		269.1557	–	–	✓	+++	+++
(15) 26.4	Substrate		271.1557	✓	✓	✓	+++	+++
OrgX								
9.4	+ 2 x O	[M+K] ⁺ = 379.1298	339.1612	–	✓	✓	–	–
10.3	+ 2 x O	[M+K] ⁺ = 379.1307	339.1615	–	✓	✓	–	–
(16) 14.2	+ 2 x O	[M+K] ⁺ = 379.1309	339.1598	–	✓	✓	++	–
(17) 16.5	+ 2 x O	[M+K] ⁺ = 379.1299	339.1595	✓	✓	✓	++	–
(18) 19.0	+ O	[M+Na] ⁺ = 363.1354	323.1658	✓	✓	✓	+	–
(19) 20.4	+ O	[M+Na] ⁺ = 363.1359	323.1650	✓	✓	✓	+	–
(20) 21.5	- 2H + O	[M+Na] ⁺ = 347.1409	321.1490	–	✓	✓	+	–
(21) 22.4	- 2H + O		321.1507	–	✓	✓	+++	+++
25.0	Substrate		307.1699	✓	✓	✓	+	+
(22) 26.1	- 2H		305.1552	–	✓	✓	+++	+++
17α-E2								
8.1	- 2H + 2 x O		301.1445	–	✓	–	–	–
9.2	+ 2 x O		303.1601	✓	✓	–	–	–
9.6	- 2H + 2 x O		301.1445	–	✓	–	–	–
12.0	+ O		287.1662	✓	✓	–	+	–
14.2	- 2H + O		285.1502	–	✓	–	+	–
15.0	+ 2 x O		303.1612	–	–	✓	+	+
15.2	+ O		287.1663	✓	✓	✓	+	+
16.0	- 2H + O		285.1507	✓	✓	✓	+	–
17.1	- 2H + O		285.1507	✓	✓	✓	+	+

Determination and identification of estrogenic compounds

T _r (min)	Modification	m/z + ion mode	m/z – ion mode	M11 _{his}	M02 _{his}	HLM	hERα	hERβ
17.7	+ O		287.1659	–	–	✓	–	–
19.1	- 2H + O		285.1507	✓	✓	–	++	–
20.1	+ O		287.1652	✓	✓	✓	++	++
20.7	+ O		287.1645	✓	✓	✓	–	–
21.9	- 2H + O		285.1506	✓	✓	✓	–	–
22.4	- 2H + O		285.1507	–	–	✓	++	–
24.1	+ O		287.1657	–	–	✓	++	–
26.1	- 2H		269.1549	✓	✓	✓	++	++
27.1	Substrate		271.1557	✓	✓	✓	++	++
Org Y								
15.8	+ 2 x O	[M+K] ⁺ = 397.1777	357.2075	✓	–	–	–	–
17.7	+ 2 x O	[M+K] ⁺ = 397.1780	357.2088	✓	✓	✓	–	–
19.0	+ 2 x O	[M+K] ⁺ = 397.1786	357.2072	✓	✓	✓	–	–
19.6	+ O	[M+K] ⁺ = 381.1830	341.2133	✓	✓	✓	–	–
20.5	+ O	[M+K] ⁺ = 381.1834	341.2131	–	–	✓	++	–
21.3	- 2H + O	[M+K] ⁺ = 379.1675	339.1967	–	–	✓	+++	+++
22.2	- 2H + O		339.1957	–	–	✓	++	++
22.4	+ O	[M+K] ⁺ = 381.1821	341.2116	✓	✓	✓	–	–
23.3	- 2H + O	[M+K] ⁺ = 379.1679	339.1950	–	–	✓	++	–
24.9	- 2H + O	[M+K] ⁺ = 379.1683	339.1966	✓	✓	✓	++	++
25.9	+ O		341.2123	✓	✓	✓	+++	+
27.4	+ 2H + O	[M+K] ⁺ = 383.1993	343.2274	✓	✓	–	–	–
27.5	+ O		341.2134	–	–	✓	++	+
29.8	- 2H	[M+H] ⁺ = 325.2183	323.2022	–	–	✓	–	–
30.0	Substrate	[M+K] ⁺ = 365.1880	325.2183	✓	✓	✓	++	++
32.1	- 4H + O		337.1812	–	–	✓	++	–

NMR analysis

The structures of the M02_{his} produced entities were mainly determined by comparing the NMR spectra of the biosynthetic produced entities with the spectra of the parent compound (NET). This part of the supplementary explains the NMR structure determination of M02-7, as an example. From mass spectrometry data, it is already known that the modification of M02-7 is NET+O.

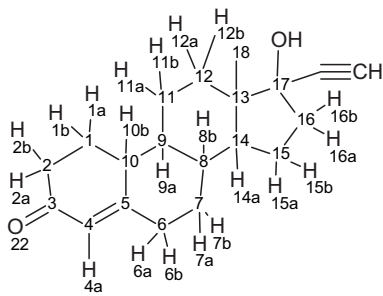
Figure S7 shows the HSQC spectra of NET (black structure and assignments) and M02-7 (red structure and assignments). The shift of the C-H correlations 15 (blue circles/arrows), and 16 (light blue circles/arrows) indicates that the hydroxyl of M02-7 is located at position 15 or 16. If the hydroxyl is located at position 15, one would expect a chemical shift change of the C14-H14 correlation towards lower field, which is not observed. So, by only using the HSQC spectra of NET and M02-7, one can easily determine the position of hydroxylation (C16).

Figure S8 shows the DQF-COSY (Double Quantum Filtered Correlation Spectroscopy) spectra of NET and M02-7. As can be noticed directly, these two spectra are highly comparable, the major difference is that H16 shifted from ~2.0 ppm to ~3.9 ppm and that both H15a and H15b shifted in frequency. Notice that H16 has just two correlations with 15a and 15b, whereas most other protons have a more complex coupling pattern. By analyzing the complete DQF-COSY spectra, it can be confirmed that the OH modification occurs at C16.

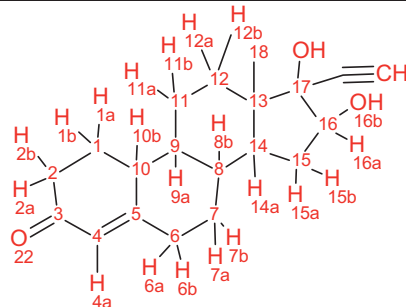
The 3D structure of M02-7 was determined by measuring a NOESY (Nuclear Overhauser Enhancement Spectroscopy) spectrum. According to the presence of a NOE correlation between H16a and H14a it allows the conclusion that the 16OH is located beta. The letters a and b used in Figure S7 and S8 refer to the belonging proton located at the alpha (a) or beta (b) position. The rest of M02-7 has the same 3D configuration as NET. For complete 3D structures of NET and M02-7, see Figure 5.

Figure S7 (next page): HSQC spectra of NET (upper) and M02-07 (lower). Large circles indicating the shift of the C15-H15 correlations, smaller circles and arrows indicating the shift of the C16-H16 correlation. Note that H4 is not represented in this figure ($\delta_{\text{H}}=5.7$ ppm, $\delta_{\text{C}}=124.0$ ppm).

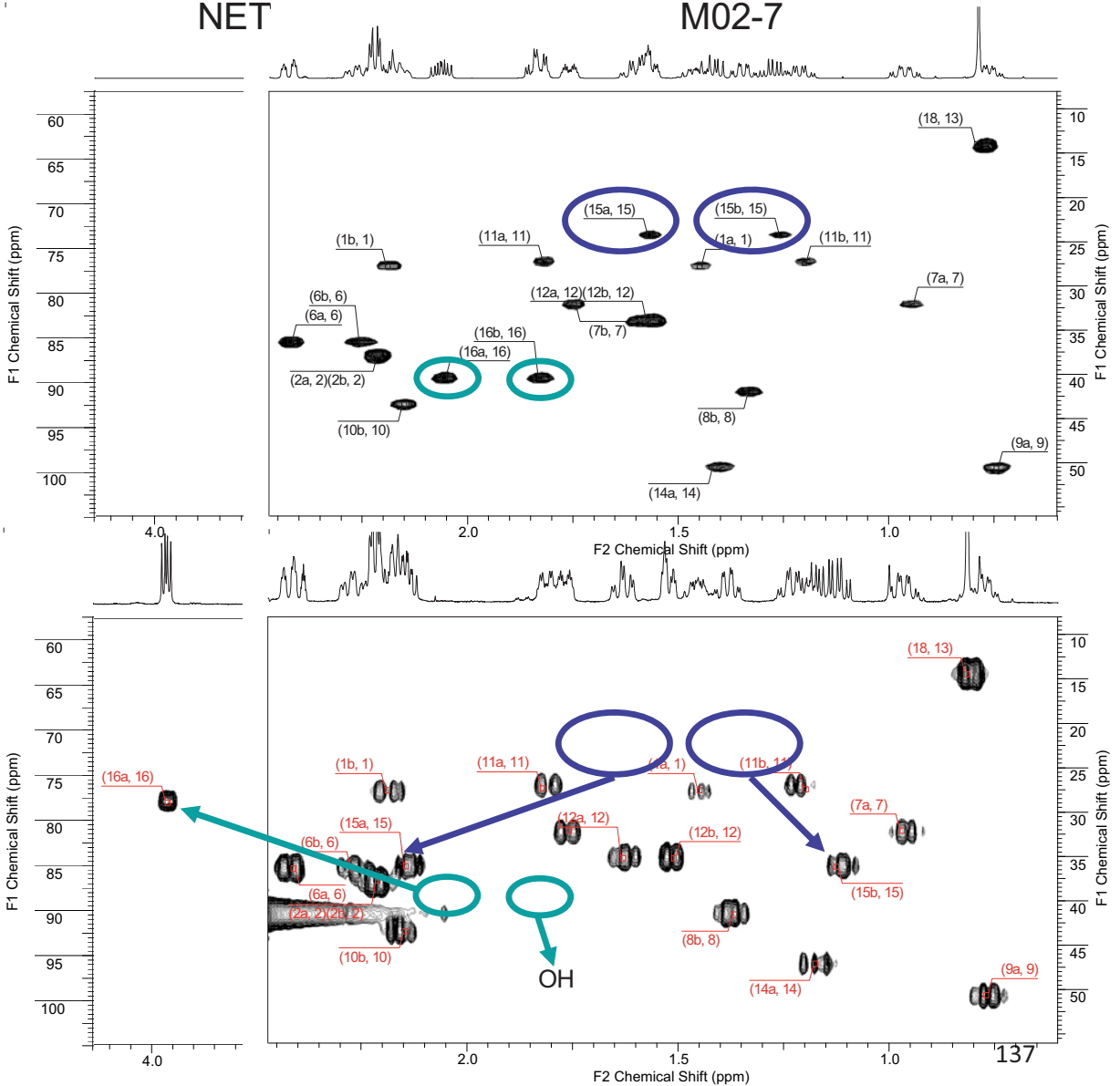
Determination and identification of estrogenic compounds



NET



M02-7



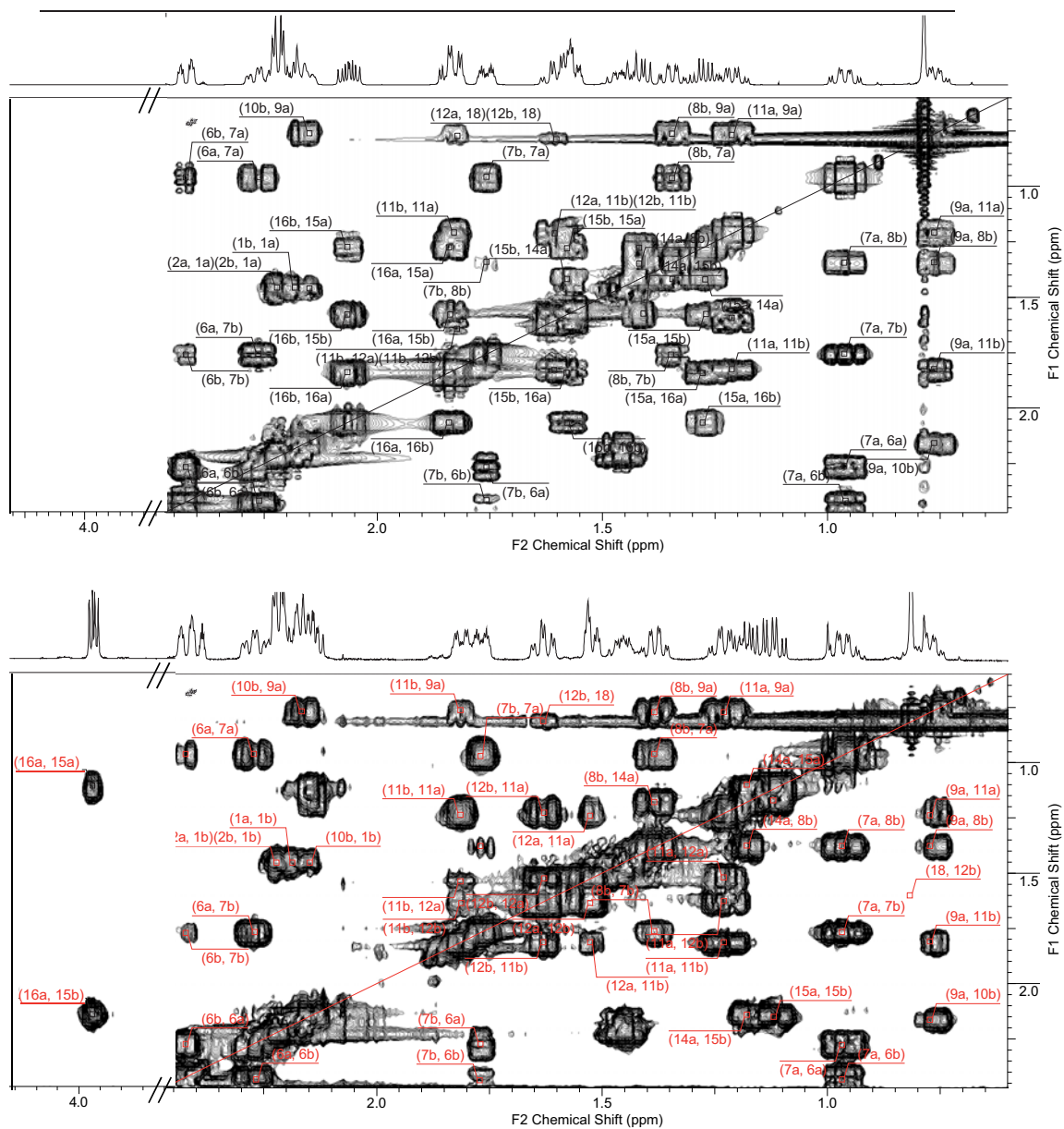


Figure S8: DQF-COSY spectra of NET (upper) and M02-07(lower). The different coupling pattern of M02-7 confirms that hydroxylation occurs at C16. Note that H4 is not represented in this figure ($\delta_{\text{H}}=5.7$ ppm).